互联网金融平台中高违约风险用户识别算法 *

阳晓慧，郭炳晖†，米志龙，郑志明

(北京航空航天大学 数学与系统科学学院 数学信息与行为教育部重点实验室，北京 100191)

摘 要：在我国金融科技不断创新的背景下，互联网金融平台中通过网络分析技术开展用户风险识别已经成为当前的热点技术发展方向。以某互联网金融平台的用户交易数据为对象，通过分析其中借贷逾期违约的传播行为，提出通过传播特征构建模型算法识别互联网金融平台的高风险用户。在构建基于阈值传播和随机传播的 SIS 模型和 SIR 模型基础上，将模型转换为可评价用户风险值的算法，并进一步与实际违约数据进行验证对比。对比结果显示在前 5%和 10%高风险群体划分条件下，算法具有较高的召回率和良好的结构关联性。

关键词：风险传播；复杂网络；互联网金融；识别算法

中图分类号：F724.6　　doi: 10.3969/j.issn.1001-3695.2017.09.0928

Identification algorithm of high breaching risk member for Internet financial platform

Yang Xiaohui, Guo Binghui†, Mi Zhilong, Zheng Zhiming

(*School of Mathematics &Systems Science*, *Beihang University*, *Beijing* 100191, *China*)

Abstract: At the background of increasingly developed financial technology, identifying of high-risk members through network analysis method on internet financial platform has become a newly hot research area. In this paper, we studied loan transaction data of an internet financial platform, and identified the high-risk members by analyzing the propagation behavior of the loan network. We established the SIS model and SIR model based on threshold propagation and random propagation, respectively. After that, we generated an algorithm to evaluate the users' risk value. Furthermore, we compared it with the actual defaulting data. In terms of the top 5% and 10% high risk group division, the results show that we can achieve high recall rate and good structural correlation with the algorithm.

Key Words: risk propagation; complex network; internet banking; identification algorithm

0 引言

随着我国经济的蓬勃发展和金融科技的不断创新，联网金融平台在金融市场中迅速发展壮大。我国的互联网金融平台在个人借贷频繁化的背景下，过互联网平台实现贷款业务的网络化运营，其中最典型的代表是 P2P 借贷平台，熊亚骅等对 P2P 借贷的信用机理、理论基础、风险来源等进行了细致探讨 [1]。为有效保障互联网金融平台在政策法规之下健康地融入现有金融体系，数据和网络分析技术构建控制系统风险模型、开展用户风险等级识别等研究已经成为当前具有现实意义的前沿热点领域。

目前已经有学者从不同角度对违约风险开展了研究工作。王书斌等人[2]对 P2P 网贷违约风险及其传染性评估的现状及方法进行了综述研究；王倩等人[3]对信用违约的风险传染模式及规律建模并给出了实证；李杰汀[4]将关联信用风险传染在复杂网络关联结构条件下进行了建模与仿真；陈庭强等人[5]系统分析在复杂网络上的信用风险传染行为演化机制；王申侃等人[6]通过信息不对称理论和线性回归分析法，研究了借款利率、借款期限、信用评级等因素同违约率之间的相关性；丁岚等人[7]通过在初级学习器、次级学习器的基础上构建集成策略模型，对 P2P 借贷中用户的违约风险进行评估和预测；唐剑琴等人[8]通过改进后的 C4.5 基本决策树模型，对信用进行评估，得到了解释力较强的结论。

在互联网金融平台中，由于信用画像基于互联网数据而非实际业务人员接触获得，用户违约风险预测大多基于对用户信用风险与用户特征之间的关联分析，而基于网络关系的可能会影响用户违约的因素还没有被挖掘出来。本文在对实际交易数据和用户网络关联结构分析的基础上，针对某互联网金融中介平台中被动违约用户行为展开研究，发现了其中的债务违约传递特征并使用复杂网络传染扩散模型构建了面向被动违约高危

基金项目：国家自然科学基金资助项目（11401017，11671025）；国家自然科学基金重大项目（11290141）

作者简介：阳晓慧（1992-），女，硕士研究生，主要研究方向为复杂网络与数据分析；郭炳晖（1982-），男，（通信作者），讲师，博士，主要研究方向为复杂系统、金融风险量化（guobinghui@buaa.edu.cn）；米志龙（1993-），男，博士研究生，主要研究方向为复杂网络与数据分析；郑志明（1953-），男，教授，博士，主要研究方向为复杂信息系统．

群体识别的模型及相关算法。

通过定义确定型和阈值型的传播模式，本文在违约免疫和非违约免疫两种情况下提出了反映上级违约造成被动违约的风险传播模型。将基于社交网络的互联网金融平台系统性风险和高危群体识别量化。该模型使用复杂网络结构参数度分布、PageRank 中心性、介数中心性[9]等作为扩散行为基本参数，通过网络传播模拟得到了单个用户违约风险传播风险评价值。经过与平台实际数据对比发现，模型在识别借贷关系中容易被他人违约影响并且容易导致他人违约的高危群体具有一定效果。

1　基于违约风险传播的识别模型[10,11]

1.1　模型思路

在对某互联网金融平台借贷数据的分析中发现，给系统带来风险的逾期违约交易中，有相当比例是因为用户在借贷链条中的上级节点违约所导致。本文针对借贷链条中的风险传播行为，提出基于被违约阈值和传播影响力的违约风险传播识别模型。用户被违约阈值$\emptyset_i$的设定依据是过去 12 个月中该用户所有借贷交易的平均值。由于在实际的金融风险控制中，用户一旦产生预期违约交易，互联网金融平台一般会根据实际情况限制该用户的借贷行为。因此，本文提出无限制规则的 QSIS 模型来描述无限制违约用户借贷行为的系统平台风险识别，而用有限制规则的 QSIR 模型来描述有限制违约用户借贷行为的系统平台风险识别。其中，用户分为未违约用户（S）、违约用户（I）和限制交易用户（R）。进一步从复杂网络的角度出发，当同时有多笔交易时考虑网络的结构对传播的影响，在度分布中心性、介数中心性和 pagerank 中心性三种结构参数的条件下考虑确定型传播和概率传播两种不同的传播模式，研究思路如图 1 所示。

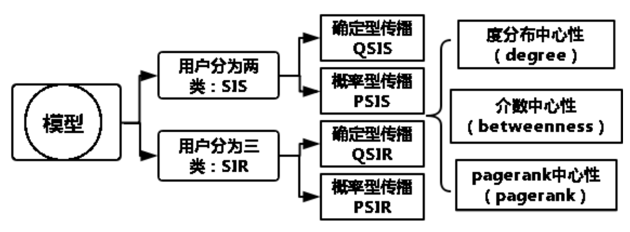


图 1　模型总体结构图

1.2　基于确定型违约风险传播的识别模型 QSIS

以实际交易数据建立网络$G(V,E)$，节点V表示所有交易用户的集合，在 QSIS 模型中，将用户分为两种类型的节点：未违约用户（S）和违约用户（I），记 $S(x)$为未违约用户的集合，$I(x)$为违约用户的集合。E表示网络中用户之间交易的集合，边E_{ij}上的权值记为$\{a_{ij}\}$，表示用户之间交易的金额，考虑到两两用户之间往往有多笔交易，在研究扩散过程时，通过重边压缩的方法处理两两用户之间的多重交易。实验中比较了平均值和最大值的模拟结果后，选取两两用户之间所有交易的最大交易金额作为该用户之间的交易金额。

记用户 i 的状态为n_i，$n_i = 1$表示违约，$n_i = 0$表示未违约；记用户之间交易E_{ij}的状态为e_{ij}，$e_{ij} = 1$表示该条用户之间的交易违约，$e_{ij} = 0$表示交易不违约。用户的被违约金额$d_j = \sum A_{ij} a_{ij} e_{ij}$,若用户被违约金额的阈值分布记为$\{\emptyset_i\}$，定义用户被传播违约的次数为用户的传播违约力大小，反映用户在该借贷网络“被违约”的强弱大小，记为$\{F_i\}$，记整个扩散过程中违约风险传播到的“被违约”用户的集合为$Lie(x)$。则以基于 SIS 确定型的阈值传播模型(QSIS)为例，整个过程可以用下述模型[12]描述：

a)初始状态，所有人都处于未违约状态（S），随机使一部分用户变为违约状态（I），即随机使一部分n_i由 0 变为 1，于是这一部分用户的某一个确定的未到期的交易E_{ij}发生违约，$e_{ij} = 1$。

b)违约的交易金额累加到债权人身上，一旦债权人被违约金额大于所能承受的阈值，即当$d_j = \sum A_{ij} a_{ij} e_{ij} > \emptyset_j$,该债权人状态由 S 变成 I，该被违约的用户将违约其持有的确定的未到期交易，违约金额累加到债权人身上。

c)判断所有用户状态不再改变，实现一次网络传播过程，新一轮传播时，用户均恢复到未违约状态（S），重新具有违约的可能，随机使一部分用户违约，开始下一轮传播。

d)反复进行上述的模拟传播过程，记录下每个用户被传染违约变为 I 状态的次数，以此定义为他们的违约传播力大小，当前后两次传播结束后被违约的用户集合$Lie(x)$相同时，传播过程结束。

1.3　基于概率型违约风险传播的识别模型 PSIS

概率型传播模型 PSIS 与确定型传播模型 QSIS 的区别在于，当一个用户违约时，不再确定性地违约某一笔交易，而是以一定的概率违约某一笔或者某几笔交易。在本文中，具体步骤与上述 QSIS 模型的步骤的在第一和第二步上有区别，即当用户 i 违约时，i 所涉及到的交易E_{ij}会以一定的概率违约，这个变化的概率与网络以上三个结构参数相关。

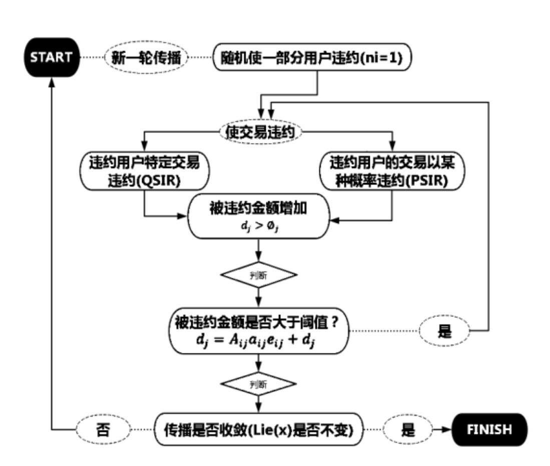


图 2　无限制规则条件下的基于违约风险传播的识别模型

1.4　基于两种风险传播的有限制模型 QSIR 和 PSIR

由于实际风险控制需要，互联网金融平台可能会对违约风险较大的用户限制其在平台中的借入借出交易。在有限制交易的条件下，本文提出有限制的确定性和概率型违约风险传播过程，并构建相应的识别模型 QSIR 和 PSIR。QSIR 与 QSIS 在于

每一轮传播中，使用随机免疫策略随机免疫一部分用户，使其不受上级节点的违约影响也不能将违约传播至下级节点。

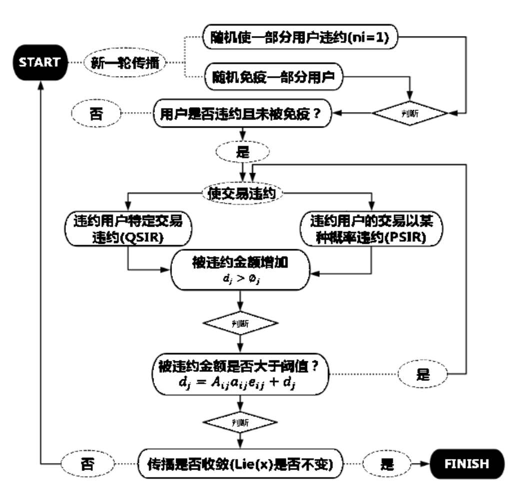


图 3　有限制规则条件下的基于违约风险传播的识别模型

2　高风险用户识别算法实现

2.1　数据来源及处理

2.1.1 原始数据

数据来自某金融借贷平台，选取其中用户 3312 个，选取 2015.5.14-2017.4.24 之间的所有交易数据，共 860999 条交易，其中违约用户有 872 个，记这 872 个用户的序列为$D(x)$；经分析后发现，由于一方债务人的违约，导致该债权人违约的用户有 213 个，将这些用户的序列记为 *L(x)*。对用户之间多个交易进行重边压缩并构建网络（图 4），并从交易网络中抽取出违约交易构成子网络（图 5）。

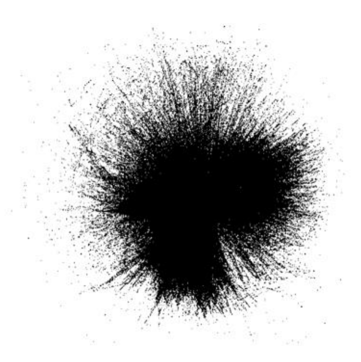


图 4　某金融借贷平台 2015.5.14-2017.4.24 用户交易结构图

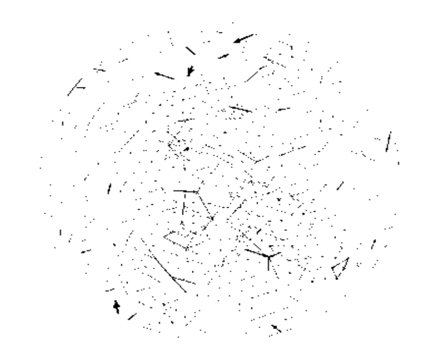


图 5　2015.5.14-2017.4.24 违约交易结构图

2.1.2 对比数据组数据

选取 2016.7.22-2017.7.22 的上述用户的交易数据作为对比数据组，同样，在 2016.7.22-2017.7.22 间涉及到的交易用户有 3312 人，共有 537146 条交易，其中违约交易有 11765 条记录，有 711 人违约他人，其中传播违约的人数有 43 人。记对比数据集为C(x)，将本文的模型用在对比数据集上，计算该数据集下的召回率，记为$R(x)$。

2.2　网络结构特征参数

本文使用到以下三个网络结构特征参数：

a)有向网络的度中心性。出度$k_i^{in} = \sum_{j=1}^{n} A_{ij}$，入度：$k_i^{out} = \sum_{j=1}^{n} A_{ij}$，其中$\{A_{ij}\}$是网络中的邻接矩阵。

b)pagerank 中心性。$x_i = \alpha \sum_j A_{ij} \frac{x_j}{k_j^{out}} + \beta$，其中$\alpha$和$\beta$是正常数。

c)介数中心性。$x_i = \sum_{st} \frac{n_{st}^i}{g_{st}}$，其中，$n_{st}^i$是从 s 到 t 经过 i 的测地路径数量，$g_{st}$为从 s 到 t 的测地路径总数。

2.3　阈值选择

2015.5.14-2017.4.24 之前选取的 3312 个用户间共有 860999 条交易记录，其中违约交易共有 16457 条，在 95%的置信区间内，所有违约交易涉及到的金额分布如图 6 所示。

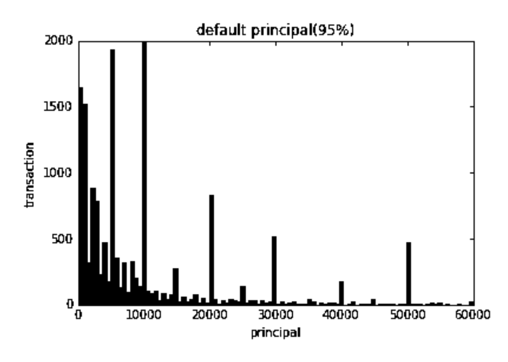


图 6　所有违约交易金额分布图（95%置信区间）

95%置信区间下的所有交易的平均金额是 10367，因此本文选取从 100 到 10000 均匀分布的 100 个阈值做实验。

2.4　实际数据验证

本文在 100-10000 间均匀地选取 100 个阈值，并分别在 QSIS、PSIS、QSIR、PSIR 等四种模型下进行违约风险扩散研究。在每一种模型下，分别分析度分布、介数分布、pagerank 分布下扩散行为的异同，在每一种扩散模式下，得到被违约的用户集合$Lie(x)$和这些用户的违约力大小分布$\{F_i\}$。比较不同扩散模型下对违约用户和因风险扩散而“被违约”的用户的预测精度，本文中通过定义召回率来刻画违约用户中因风险扩散而违约的用户比例的预测精度，并以此作为评价模型好坏的标准。召回率的值是预测正确的实际由于扩散而违约的人与预测正确的的实际违约人数之比，即

$$\frac{Lie(x) \cap L(x)}{Lie(x) \cap D(x)}$$

召回率的定义反映了违约用户中经扩散违约的用户的预测比例，反映了局部扩散违约用户预测的准确率大小，是衡量对传播违约用户预测精度的有利依据。分别对比高危群体前 10%和 5%的用户在不同模型下的扩散结果。

在 PSIS 模型和 QSIS 模型下，根据构建的交易网络，用户分为两类：未违约用户（S）和违约用户（I），每次传播随机选择600个用户违约，看违约风险在网络中的传播情况。在PSIR模型和 QSIR 模型下，根据构建的交易网络，将用户分为三类：未违约用户（S）和违约用户（I），和拉黑用户（R，也叫限制交易用户），同样每次传播随机选择 600 个用户违约，看违约风险在网络中的传播情况。当风险扩散过程中“被违约”的用户集合$Lie(x)$不再发生变化时，扩散过程收敛，停止传播。计算出$Lie(x)$中用户F_i排序后的序列，选取高危性前 10%的用户，此时$Lie(x)$指的是影响力大小$\{F_i\}$从高到低排序的前 332 个用户。比无限制借贷模式下，两种传播方式下的预测精度。对比 QSIS 和 PSIS 模型下前 10%的高危用户的相关结论如表 1 所示。

表 1　前 10%用户 QSIS 模型分析

QSIS	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	110	49	44.55%	22.53%
pagerank	10	10	100%	100%
betweenness	110	46	41.82%	19.72%

表 2　前 10%用户 PSIS 模型分析

PSIS	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	137	71	51.82%	10.17%
pagerank	91	34	37.36%	18.42%
betweenness	89	30	33.73%	7.69%

对原始数据的结论进行分析，对比表 1 和表 2 可以看出，无限制借贷模式，PSIS 模型下以度分布传播时，对违约用户中的传播违约用户预测的准确率最高。整体上从对违约用户中的传播用户的预测比例来看，QSIS 模型的预测效果强于 PSIS 模型。而在 QSIS 模型中以结构参数 pagerank 预测的高危用户几乎全是传播违约的用户，这一结论在对比数据集上也成立。

同样，对于有限制借贷模式，计算$Lie(x)$中用户F_i排序后的序列，选取高危性前 10%的用户，此时$Lie(x)$指的是影响力大小$\{F_i\}$从高到低排序的前 332 个用户，并对比两种传播方式下的预测精度。对比 QSIR 和 PSIR 的相关结论如下表所示：

表 3　前 10%用户 PSIR 模型分析

PSIR	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
Degree	134	75	55.97%	18.30%
pagerank	140	74	52.86%	7.69%
betweenness	140	73	51.80%	23.81%

表 4　前 10%用户 QSIR 模型分析

QSIR	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	110	49	44.76%	22.54%
pagerank	115	56	48.67%	100%
betweenness	109	48	44.04%	19.72%

对比表 3 和 4 可以看出，PSIR 模型下以度分布传播时，对违约用户中的传播违约用户预测的准确率最高。在原始数据集上，无论是以哪一种结构参数下的扩散模型，以概率传播时的 SIR 模型对违约用户中的传播违约用户的准确率比确定传播时的 SIR 模型高，整体来说，对比数据集下预测的准确率偏低。

而在实际中，人们往往关注风险特别大的高危群体，在本文中，为了进一步评价模型对高危群体的预测准确度，计算出用户传播影响力F_i排序后的序列，并取高危性前 5%的用户序列，记为新的$Lie(x)$，此时$Lie(x)$指的是影响力大小$\{F_i\}$从高到低排序的前 166 个用户，在无限制借贷行为模式下，对比两种传播方式下的预测精度。对比 QSIS 和 PSIS 模型前 5%的高危用户的相关结论如表 5、6 所示.

表 5　前 5%用户 QSIS 模型分析

QSIS	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	100	30	30.00%	17.19%
pagerank	10	10	100%	100%
betweenness	61	31	50.85%	24.23%

表 6　前 5%用户 PSIS 模型分析

PSIS	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	131	43	32.82%	12.02%
pagerank	84	19	22.64%	13.15%
betweenness	41	19	46.34%	14.57%

对比表 1~5 发现，PSIS 模型下对前 10%的高危群体中传播违约用户的预测精度高于对前 5%的高危群体中传播违约用户的预测精度，这个结论对 QSIS 模型依然成立，通过对比表 2 和 6 就可以得知。这说明无限制借贷模式对于借贷网络中传播风险更高的用户预测的准确度会下降。观察表 5 可以发现一个跟表 1 相同的现象，结构参数 pagerank 下预测高危用户的数量虽然不多，但是找到的全是传播违约的用户，精度达到了 100%，同样，这一结论在对比数据集上也成立。

同样，对于有限制借贷模式，计算出$Lie(x)$中用户F_i排序后的序列，选取高危性前 5%的用户，此时$Lie(x)$指的是影响力大小$\{F_i\}$从高到低排序的前 166 个用户，并对比两种传播方式下的预测精度。对比 QSIR 和 PSIR 模型前 5%的高危用户的相关结论如表 7、8 所示。

表 7　前 5%用户 PSIR 模型分析

PSIR	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	128	45	35.16%	23.81%
pagerank	140	50	35.77%	24.39%
betweenness	73	40	54.93%	16.67%

表 8　前 5%用户 QSIR 模型分析

QSIR	$Lie(x) \cap D(x)$	$Lie(x) \cap L(x)$	召回率	召回率 $R(x)$
degree	103	30	29.13%	17.19%
pagerank	100	34	33.98%	100%
betweenness	61	30	49.15%	24.32%

对比表 3 和 7 后，对原始数据集的模拟结果进行分析，发现整体上 PSIR 模型下对前 10%的高危群体中传播违约用户的预测精度高于对前 5%的高危群体中传播违约用户的预测精度，但是介数传播下的结果例外；这个结论对 QSIS 模型依然成立，通过对比表 4 和 8 就可以得知。这说明有限制借贷模式对于网络传播风险中风险更高的用户预测的准确度会下降，但是介数传播下对更高危的传播违约用户的预测精度更高。

比较两个数据集下的结果，发现对比数据集整体上预测效果更差，但是 pagerank 的集中预测性能更好。

在对比上述表 1~8 后发现，对原始数据集的模拟结果进行分析，总的来说，采用随机免疫策略后的模型（如 PSIR 和 QSIR）对于借贷网络中传播违约的高危用户的预测精度更好；对于原始数据集来说，这四种模型下对于前 10%的高危用户的预测精度高于对前 5%高危用户的预测精度；本文中除 QSIS 模型外，其他三种模型下对于前 5%的高危传播风险用户的预测，介数中心性均取得比其他两种结构参数更好的结果。而对前 10%的高危传播风险用户的预测中，除了 PSIS 模型外，其他三种模型下 PageRank 中心性扩散下的效果均比其他两种结构参数好。

对比两种数据集下的结果，发现整体而言，对比数据集的预测效果不及原始数据集好。由于对比数据是一年的数据，而原始数据集有接近两年的数据。某种程度上说明，本文中的模型对于较长时间数据预测效果更好，而对于较短时间数据集的预测，虽然整体上预测精度有所下降，但结构参数 pagerank 下预测的不多的高危用户都是传播违约用户，预测精度达到了 100%，而且这种现象出现 比例高于对原始数据集进行分析出现的比例，这说明短期数据集对节点的传播违约集中性的预测效果更好。

2.5　模型对比分析

上述对结果的分析是从准确性的角度出发的，接下来观察三种结构参数下四种传播模式 QSIS、QSIR、PSIS、PSIR 下。选取阈值 5000 下的具体用户的传播影响力F_i的分布，归一化后进行作图。原始数据集下的计算结果如图 7.1，图 8.1，图 9.1，相应地，对比数据集下的计算结果如图 7.2、8.2、9.2 所示。

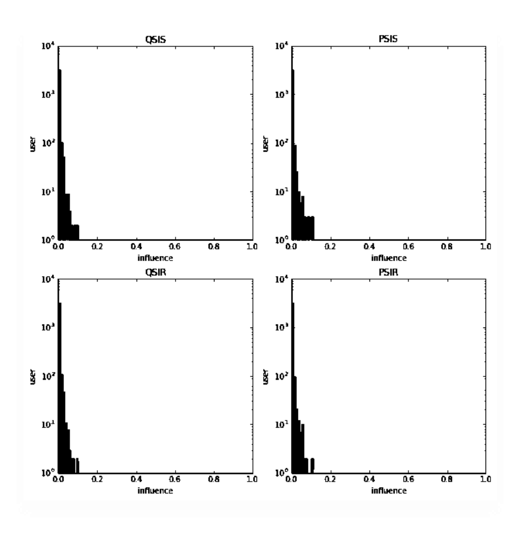


图 7　原始数据集度分布传播下四种模型F_i分布图

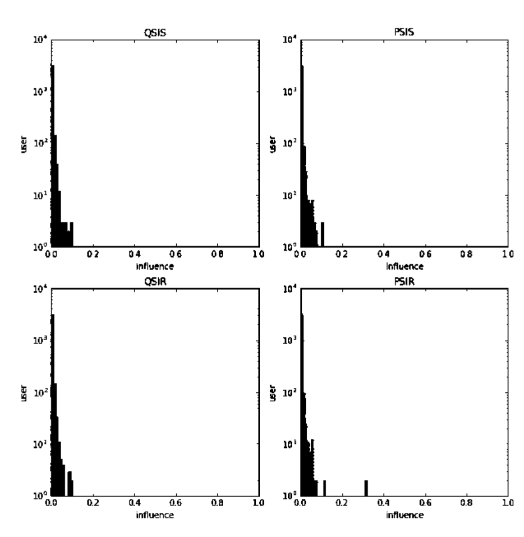


图 7　对比数据集中度分布传播下四种模型F_i分布图

图 7.1 中四张分图分别是在原始数据集上，以结构参数度分布传播时，采用扩散模型 QSIS、PSIS、QSIR、PSIR 模拟下的用户违约传播影响力F_i的分布图。而图 7.2 是在对比数据集上的实验结果。

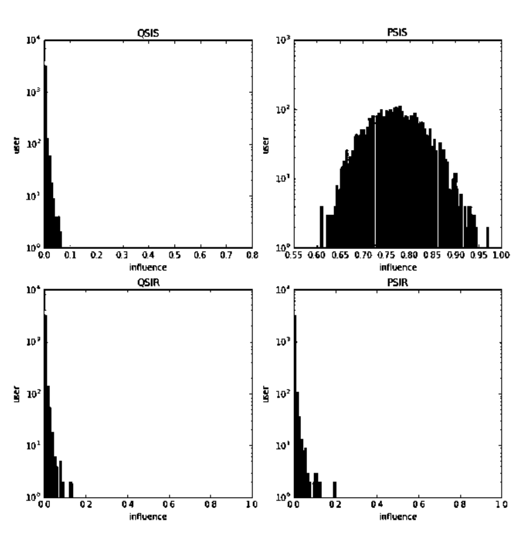


图 8　原始数据集介数分布下四种F_i分布图

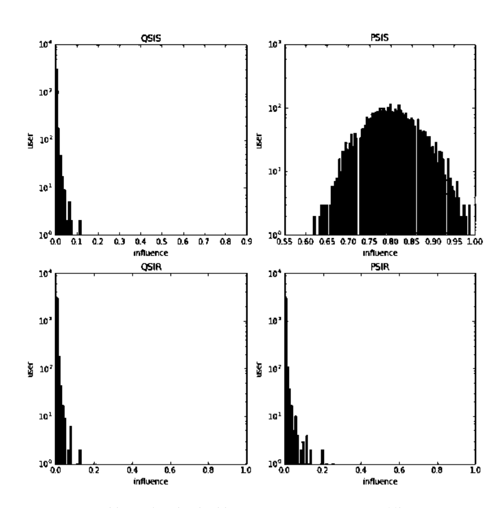


图 8　对比数据集中介数分布传播下四种模型F_i分布图

图 8.1 中四张分图分别是在原始数据集上，以结构参数——介数中心性传播时，采用扩散模型 QSIS、PSIS、QSIR、PSIR 模拟下的用户违约传播影响力F_i的分布图。而图 8.2 是在对比数据集上的实验结果。

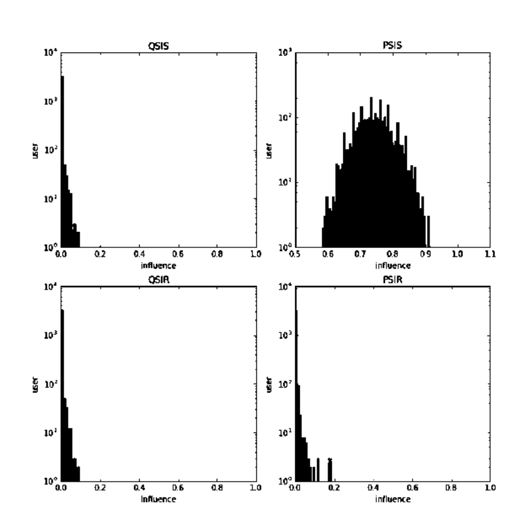


图 9　原始数据集 pagerank 传播下四种模型F_i分布图

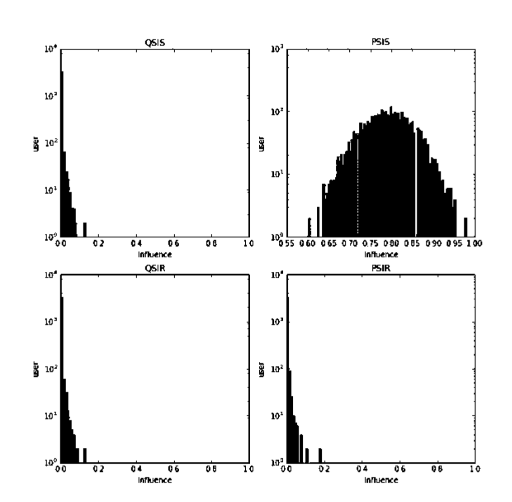


图 9　对比数据集 pagerank 传播下四种模型F_i分布图

图 9.1 中四张分图分别是在原始数据集上，以结构参数 pagerank 中心性传播时，采用扩散模型 QSIS、PSIS、QSIR、PSIR 模拟下的用户违约传播影响力F_i的分布图。而图 9.2 是在对比数据集上的实验结果。

比较上述三种网络结构参数下的图会发现，同一网络结构参数下，不同扩散模式下用户的传播违约力分布差异不大，基本上服从指数分布。而 PSIS 模型中介数和 *pagerank* 传播下的图像相比其他扩散模型有显著区别，其他模型中用户的影响力分布符合指数分布，而在这两种模式下用户的影响力分布接近于正态分布。通过对比原始数据集和对比数据集下的结果，发现相同传播模式下F_i的分布几乎相同，一定程度上说明了模型对于不同的数据集具有一定的鲁棒性。而相同扩散模式下，三种网络结构参数下扩散的F_i的分布却十分接近。

上述图表对比了阈值 5000 下，不同结构参数和不同传播模式下的F_i分布图，为了进一步观察不同阈值下的F_i分布，在 PSIS 模型下以介数中心性传播时为例，在 600 到 9600 之间等距选取 10 个阈值，其违约影响力大小分布见下图：

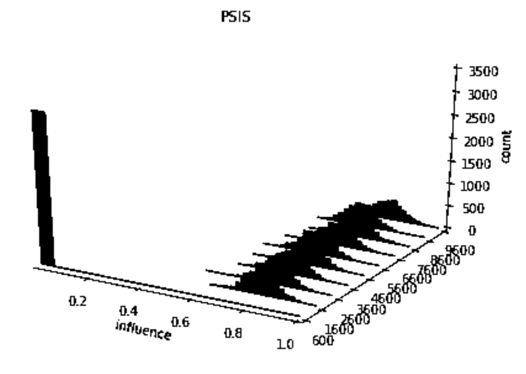


图 10　不同阈值条件下F_i的分布图

从图 10 可以看出，阈值为 600 时，出现影响力都很大的情况，这可以理解为由于债权人能承受的被违约金额太低了，此时每一次债务人违约，几乎都会导致相应的债权人违约，而当阈值为 1600 到 9600 之间时，用户的传播违约影响力大小近似呈现正态分布，此时用户可以被有效区分。

3　结束语

本文以某互联网金融平台 2015.5.14-2017.4.24 时间段内的实际交易数据为对象作为原始数据集，通过选择该平台 2016.722-2017.7.22 的相应交易作为对比数据集。通过分析违约交易传播的行为特征提出相应的计算模型以及高风险用户识别算法，并用在两个数据集上。将算法结果与实际数据比对后发现该算法具有较好预测效果，对互联网金融平台系统性风险预测及防范有广泛应用价值。

参考文献：

[1] 熊亚骅，熊一鹏，李婷．互联网金融网贷客户违约风险研究进展 [J]. 金融经济, 2015 (12): 70-7.

[2] 王书城，谭中明，陈艺云. P2P 网贷违约风险及其传染性评估综述 [J]. 武汉金融, 2017 (6): 40-44.

[3] 王倩, Hartmannwendels T. 信用违约风险传染建模 [J]. 金融研究, 2008 (10): 162-173.

[4] 李杰汀. 基于复杂网络的关联信用风险传染建模与仿真研究 [D]. 成都：电子科技大学, 2015.

[5] 陈庭强，何建敏. 基于复杂网络的信用风险传染模型研究 [J]. 中国管理科学, 2014, 28 (11): 111-117.

[6] 王申侃. P2P 网贷平台应对借款违约的风险管理研究 [D]. 北京：对外经济贸易大学, 2016.

[7] 丁岚，骆品亮. 基于 Stacking 集成策略的 P2P 网贷违约风险预警研究 [J]. 投资研究, 2017, 36 (4): 41-54.

[8] 唐剑琴. 基于决策树算法的P2P网贷借款人违约风险度量研究 [D]. 长沙：湖南师范大学, 2016.

[9] Zhou Jie, Liu Zonghua. Epidemic spreading in complex networks [J]. Frontiers of Physics in China, 2008, 3 (3) .

[10] Dorogovtsev S N, Mendes J F F. Evolution of networks [J]. Adv. Phys. 2002, 51: 1079–1187.

[11] Newman M E J. Spread of epidemic disease on networks [J]. Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) , 2002, 66.

Zhang Zike, Liu Chuang, Zhan Xiuxiu, et al. Dynamics of information diffusion and its applications on complex networks [J]. Physics Reports, 2016, 651.